

Голові разової спеціалізованої вченої ради
Національного університету «Львівська політехніка»
д.т.н., професору Яковині Віталію Степановичу

РЕЦЕНЗІЯ

кандидата технічних наук, доцента, доцента
кафедри систем штучного інтелекту

Національного університету «Львівська політехніка»

Вовк Олени Борисівни

на дисертаційну роботу Михайлишина Владислава Юрійовича

*«Система підтримки та прийняття рішень аналізу рекрутингової діяльності
на основі великих даних»*

подану до захисту на здобуття наукового ступеня **доктора філософії**

з галузі знань 12 «Інформаційні технології»

та спеціальність 122 «Комп'ютерні науки»

1. Загальна характеристика роботи

Дисертація Михайлишина Владислава Юрійовича на тему «Система підтримки та прийняття рішень аналізу рекрутингової діяльності на основі великих даних» є кваліфікаційною науковою працею на здобуття наукового ступеня доктора філософії за спеціальністю «Комп'ютерні науки». У даній роботі представлено комплексне дослідження сучасних методів обробки великих даних, кластеризації, аналізу природної мови та зворотного зв'язку, що дозволяє значно покращити процеси рекрутингу. Дисертація складається з чотирьох основних розділів, кожен з яких детально розкриває різні аспекти дослідження.

У першому розділі «Аналіз методів обробки великих даних рекрутингової сфери» автор досліджує методи, які використовуються для обробки великих даних у рекрутингу, з особливим акцентом на гетерогенність даних. Детально розглядаються характеристики, пов'язані з рекрутинговими даними, а також різні методи обробки, такі як вибірка, трансформація, зменшення шуму та вилучення ознак.

Другий розділ "Розробка алгоритму обробки даних з використанням методів аналізу природної мови та кластеризації" присвячено аналізу різних методів кластеризації даних, таких як K-means, K-medians, DBSCAN, C-means, та їх тестуванню на досліджуваному наборі даних. Оцінювання результатів роботи методів проводилося за допомогою індексів Силуета та Данна. Розроблено власний метод паралельної кластеризації даних з використанням ансамблю існуючих методів та вибору кінцевого кластера для визначених точок за допомогою голосування.

Третій розділ «Розробка методу побудови системи підтримки прийняття рішень та використання зворотнього зв'язку для покращення рекомендацій» аналізує методи регресії, такі як дерева рішень, випадковий ліс, XGBoost, і виконує оцінку їх роботи за допомогою статистичних показників, вибираючи найкращий метод для застосування в системі.

Четвертий розділ «Розробка архітектури і апробація результатів» описує розроблення системи збору, опрацювання, аналізування та оцінювання відгуків кандидатів для покращення процесів рекрутингу. Використано поглиблений аналіз даних зворотного зв'язку, зокрема оцінку їхньої якості та корисності, аналіз тенденцій, визначення ключових сфер для вдосконалення та виявлення трендів. Представлено розроблену архітектуру системи, обґрунтовано обрані технології та технічні рішення під час реалізації системи.

Таким чином, дисертація Михайлишина Владислава Юрійовича демонструє високий рівень теоретичної підготовки та практичної реалізації, що відображено у вигляді розробленої системи підтримки прийняття рішень, яка вже впроваджена у навчальний процес та використовується в комерційній діяльності. Загальна структура та логіка роботи сприяють чіткому розумінню всіх аспектів дослідження, а використання сучасних наукових методів забезпечує високу надійність отриманих результатів.

2. Оцінка актуальності теми дисертації

Актуальність теми дисертаційного дослідження є безсумнівною, оскільки управління людськими ресурсами, зокрема рекрутинг, є однією з ключових функцій будь-якої організації. Сучасний ринок праці вимагає нових підходів до підбору та оцінки кандидатів, особливо в умовах значного обсягу даних, які потрібно обробляти для прийняття ефективних рішень. Великі дані відкривають нові можливості для оптимізації рекрутингових процесів, і дана дисертація пропонує конкретні методи та засоби для реалізації цих можливостей.

Значення використання великих даних у рекрутингу важко переоцінити. У сучасному світі, де інформація накопичується у величезних обсягах і різноманітних форматах, традиційні методи обробки даних виявляються недостатньо ефективними. Це особливо актуально для процесів рекрутингу, де потрібно швидко й точно оцінювати велику кількість кандидатів, їхні резюме та інші дані. Застосування сучасних технологій, таких як машинне навчання та аналіз природної мови, дозволяє значно підвищити точність і швидкість цих процесів.

Останні глобальні події, такі як пандемія Covid-19, економічні кризи та збройні конфлікти, призвели до значних змін на ринку праці. Наприклад, багато компаній перейшли на віддалену роботу, що змінило вимоги до кандидатів та способи їх пошуку.

Також зросла конкуренція за таланти, особливо в технологічних сферах, що вимагає нових підходів до рекрутингу. Система підтримки прийняття рішень на основі великих даних, розроблена автором, допомагає вирішити ці проблеми, забезпечуючи більш точний і об'єктивний підхід до відбору кандидатів.

Крім того, важливо враховувати етичні та правові аспекти збору та обробки даних. Захист персональної інформації та дотримання конфіденційності є критичними у сучасному рекрутингу. Система, розроблена автором, також враховує ці аспекти, забезпечуючи відповідність правовим нормам і підвищуючи довіру між кандидатами та роботодавцями.

Таким чином, актуальність теми дослідження обумовлена необхідністю вдосконалення процесів рекрутингу за допомогою сучасних технологій обробки великих даних. Розроблена система підтримки прийняття рішень є відповіддю на сучасні виклики ринку праці, забезпечуючи ефективність, точність і об'єктивність у відборі кандидатів. Ця робота має значний потенціал для подальшого розвитку та впровадження у різних сферах бізнесу.

3. Наукова цінність дисертаційного дослідження

Наукова цінність дисертаційної роботи Михайлишина В.Ю. полягає у розробленні нових методів та алгоритмів для підтримки прийняття рішень у рекрутинговій діяльності. Одним з основних досягнень є розробка методу кластеризації неоднорідних напівструктурованих даних на основі ансамблю методів кластеризації. Цей метод відрізняється від існуючих застосуванням попередньої зваженої токенізації та пакетним розподіленим опрацюванням даних, що дозволяє підвищити якість кластеризації.

Дослідження також включає вдосконалення методу побудови систем підтримки прийняття рішень на основі даних з використанням алгоритму градієнтного бустингу. Це дозволяє усунути проблему перенавчання та забезпечити більш точну оцінку кандидатів. Крім того, автором було розроблено комбінований метод зменшення розмірності даних рекрутингової діяльності, що дозволяє збільшити точність та швидкість подальшого опрацювання таких даних.

Оригінальність та новизна дослідження полягає в інтеграції сучасних методів машинного навчання та аналізу природної мови для вирішення практичних задач рекрутингу. Автор досліджує та впроваджує різні підходи до обробки великих даних, зокрема методи токенізації, кластеризації, регресії та аналізу зворотного зв'язку.

Використання цих методів дозволяє створити ефективну систему підтримки прийняття рішень, яка може бути використана у різних сферах бізнесу для оптимізації процесів підбору персоналу.

Наукові результати дисертації опубліковані в авторитетних наукових виданнях, зокрема у виданнях, що індексуються у міжнародних науково-метричних базах даних. Це свідчить про високий рівень дослідження та визнання його результатів науковою спільнотою. Крім того, результати роботи були представлені на міжнародних конференціях, де отримали позитивні відгуки від фахівців у галузі комп'ютерних наук та рекрутингу.

Таким чином, наукова цінність дисертаційного дослідження Михайлишина Владислава Юрійовича полягає у розробленні нових, ефективних методів обробки великих даних для підтримки прийняття рішень у рекрутинговій діяльності. Ці методи можуть бути використані для подальших досліджень та впроваджень у різних галузях, що робить роботу автора важливим внеском у розвиток сучасних технологій та наукових підходів до управління людськими ресурсами.

4. Оцінка практичного значення результатів роботи

Практична цінність дисертації Михайлишина Владислава Юрійовича є високою, оскільки розроблені методи та алгоритми були успішно впроваджені у реальні системи підтримки прийняття рішень, які використовуються в рекрутинговій діяльності. Розроблені методи та алгоритми реалізовані у вигляді програмного забезпечення, що дозволяє автоматизувати багато аспектів процесу підбору персоналу. Це значно зменшує витрати часу та ресурсів, одночасно підвищуючи ефективність та точність відбору кандидатів.

Впровадження розроблених методів у навчальний процес Національного університету "Львівська політехніка" при викладанні дисциплін «Проектування інформаційних систем» та «Об'єктно-орієнтоване програмування» є підтвердженням їх практичної значимості. Це дозволяє студентам отримувати сучасні знання та навички, необхідні для роботи з великими даними та системами підтримки прийняття рішень у рекрутингу.

Крім того, результати дисертації були використані для розробки системи підтримки та прийняття рішень у рекрутингу в ТзОВ «Палетний сервіс». Це підтверджує ефективність розроблених методів у реальних умовах бізнесу. Система дозволяє автоматизувати процеси підбору персоналу, покращуючи якість прийняття рішень та зменшуючи час, необхідний для оцінки кандидатів.

Одним з основних практичних результатів роботи є розробка архітектури системи з орієнтацією на веб-застосунок. Для застосування алгоритмів у режимі реального часу автор запропонував використовувати лямбда-архітектуру, що забезпечує масштабованість пам'яті та зменшує навантаження на сервер у 2 рази. Це дозволяє системі ефективно обробляти великі обсяги даних та надавати точні результати в короткі терміни.

Практичне значення дисертації також полягає у можливості використання розроблених методів та алгоритмів у різних сферах бізнесу, де необхідно обробляти великі обсяги даних для прийняття ефективних рішень. Це може бути корисно не тільки у рекрутингу, але й у маркетингу, фінансах, управлінні проектами та інших сферах, де важлива швидка та точна обробка даних.

Таким чином, практичне значення дисертації Михайлишина В.Ю. полягає у можливості широкого застосування розроблених методів у різних галузях бізнесу, що дозволяє значно підвищити ефективність процесів обробки даних та прийняття рішень. Це робить роботу автора важливим внеском у розвиток сучасних технологій управління людськими ресурсами та інших сфер діяльності.

5. Оцінка достовірності та обґрунтованості основних положень і висновків дисертації

Достовірність та обґрунтованість основних положень і висновків дисертації Михайлишина Владислава Юрійовича підтверджуються використанням сучасних наукових методів та інструментів, які забезпечують високу точність та надійність отриманих результатів. У роботі застосовано широкий спектр статистичних методів, методів машинного навчання, регресійних моделей та методів природної мовної обробки, що дозволяє забезпечити комплексний підхід до вирішення поставлених задач.

Результати дослідження підтверджені експериментальними даними, отриманими на реальних наборах даних, що використовуються у рекрутинговій діяльності. Це забезпечує високу надійність та практичну значимість розроблених методів та алгоритмів. Крім того, наукові результати дисертації опубліковані в авторитетних наукових виданнях та апробовані на міжнародних конференціях, що додатково підтверджує їх достовірність.

Автор дисертації провів ретельний аналіз сучасних методів обробки великих даних, кластеризації та аналізу природної мови, що дозволило розробити нові, ефективні методи для підтримки прийняття рішень у рекрутинговій діяльності. Кожен з методів був детально проаналізований та протестований на реальних даних, що дозволило автору зробити обґрунтовані висновки щодо їх ефективності та можливостей застосування.

Одним з важливих аспектів роботи є застосування методу градієнтного бустингу для побудови систем підтримки прийняття рішень. Цей метод дозволяє уникнути перенавчання та забезпечити високу точність оцінки кандидатів. Результати, отримані за допомогою цього методу, підтверджують його ефективність та надійність.

Автор також ретельно підходив до оцінки методів кластеризації, зокрема таких як K-means, K-medians, DBSCAN, C-means. Проведено оцінку їх роботи за допомогою індексів Силуета та Данна, що дозволило зробити обґрунтовані висновки щодо вибору найефективніших методів для конкретних задач. Випробування на реальних наборах даних показали, що розроблений метод паралельної кластеризації з використанням ансамблю існуючих методів забезпечує високу якість кластеризації та стабільність результатів.

Загалом, оцінка достовірності та обґрунтованості основних положень і висновків дисертації Михайлишина Владислава Юрійовича свідчить про високий рівень проведеного дослідження та надійність отриманих результатів. Використання сучасних наукових методів та інструментів, підтвердження результатів експериментальними даними та їх апробація на міжнародних конференціях забезпечують високу надійність і практичну значимість роботи.

6. Оцінка змісту й оформлення дисертації

Дисертація Михайлишина Владислава Юрійовича відзначається логічною структурою, послідовністю викладення матеріалу та високою якістю оформлення. Робота складається з чотирьох основних розділів, які послідовно розкривають всі аспекти досліджуваної теми, починаючи з аналізу методів обробки великих даних та закінчуючи апробацією розробленої архітектури системи. Така структура дозволяє читачу легко слідувати за думкою автора та розуміти логіку дослідження.

Текст дисертації написано на високому науковому рівні, з використанням сучасної термінології та посилань на актуальні наукові джерела. Автор демонструє глибоке знання теми та вміння застосовувати наукові методи для вирішення конкретних задач. У роботі представлено велику кількість ілюстрацій, графіків та таблиць, які доповнюють текст та сприяють кращому розумінню матеріалу.

Оформлення дисертації відповідає вимогам до наукових праць, що підтверджується наявністю анотацій українською та англійською мовами, списку використаних джерел та додатків. Автор ретельно підходив до оформлення кожного розділу, що свідчить про його відповідальність та професіоналізм.

Зміст дисертації охоплює всі основні аспекти досліджуваної теми, починаючи з теоретичного аналізу методів обробки великих даних і закінчуючи практичними результатами та їх апробацією. Кожен розділ логічно пов'язаний з попереднім, що створює цілісну картину дослідження. Вступ та висновки дисертації чітко визначають мету, завдання та основні результати роботи, що дозволяє легко оцінити її значимість та внесок у науку.

Особливо варто відзначити чіткість та ясність викладення матеріалу. Автор використовує зрозумілу та логічну мову, що робить роботу доступною для широкого кола читачів, включаючи як науковців, так і практиків у галузі рекрутингу та обробки даних. Використання сучасної термінології та пояснення складних наукових концепцій на зрозумілому рівні свідчать про високий професійний рівень автора.

Таким чином, дисертація Михайлишина Владислава Юрійовича відзначається високою якістю змісту та оформлення, що відповідає вимогам до наукових праць. Логічна структура, послідовність викладення матеріалу, використання сучасних наукових методів та ретельне оформлення свідчать про високий рівень роботи та професіоналізм автора.

7. Зауваження до дисертаційної роботи

1. В розділі 1 не вказано з яких джерел були отримані наведені приклади даних резюме.

2. В розділі 2 результати кластеризації кожного окремого методу подано лише графічно. Доцільним буде показати також певні статистичні показники для визначених кластерів.

3. В розділі 3 показано результати роботи різних методів регресії, проте недостатньо інформації про їх тренування, варто показати, наприклад, функцію втрат.

4. В розділі 4 варто додати аргументацію вибору AWS платформи для розробки програмного забезпечення системи.

5. У висновках до розділу 4 відсутня інформація про застосування паралельної роботи скраперів для пришвидшення збору даних та оптимізації витрат.

8. Загальна оцінка дисертаційної роботи

Загалом, дисертація Михайлишина Владислава Юрійовича є високоякісною науковою працею, яка зробила значний внесок у розвиток методів підтримки прийняття рішень у рекрутинговій діяльності на основі великих даних. Робота відзначається високим рівнем теоретичної підготовки, новизною отриманих наукових результатів та їх практичною значимістю. Розроблені методи та алгоритми забезпечують ефективну обробку великих обсягів даних, що дозволяє значно підвищити точність та ефективність процесів рекрутингу.

Дисертація демонструє глибоке розуміння автором сучасних методів обробки даних та їх застосування для вирішення практичних задач. Розроблені методи кластеризації, регресії та аналізу природної мови є новими та ефективними інструментами, які можуть бути використані у різних сферах бізнесу. Це свідчить про високий науковий рівень роботи та її значимість для практичного застосування.

Автор ретельно підходив до дослідження, використовуючи сучасні наукові методи та інструменти, що забезпечують високу точність та надійність отриманих результатів. Результати дослідження підтвержені експериментальними даними, отриманими на реальних наборах даних, що використовуються у рекрутинговій діяльності. Це забезпечує високу надійність та практичну значимість розроблених методів та алгоритмів.

Крім того, дисертація відзначається високою якістю оформлення та логічною структурою, що робить її доступною для широкого кола читачів. Автор використовує зрозумілу та логічну мову, що робить роботу доступною для науковців, практиків у галузі рекрутингу та обробки даних, а також для студентів. Використання сучасної термінології та пояснення складних наукових концепцій на зрозумілому рівні свідчать про високий професійний рівень автора.

Загалом, дисертація Михайлишина Владислава Юрійовича є важливим внеском у розвиток сучасних технологій обробки даних та підтримки прийняття рішень у рекрутинговій діяльності.

Зміст дисертаційної роботи, отримані основні наукові положення та висновки відповідають паспорту спеціальності 122 «Комп'ютерні науки».

Вказані у п. 7 цього відгуку зауваження щодо представленого дослідження не знижують вагомість отриманих у роботі наукових та практичних результатів і не змінюють її позитивної оцінки.

Робота відповідає вимогам «Порядку присудження ступеня доктора філософії», затвердженого постановою Кабінету Міністрів України від 12 січня 2022 р. № 44, а її автор Михайлишин Владислав Юрійович заслуговує на присудження наукового ступеня доктора філософії за спеціальністю 122 «Комп'ютерні науки».

Рецензент – кандидат технічних наук,
доцент кафедри Систем штучного інтелекту,
Національний Університет «Львівська Політехніка»

Олена ВОВК

«Підпис Вовк О.Б. засвідчую»:

Вчений секретар

Національного університету «Львівська політехніка»,

к.т.н., доц.



Роман БРИЛИНСЬКИЙ